Vibrational Spectra of Nucleic Acid Bases and Their Watson–Crick Pair Complexes

R. SANTAMARIA,[1]* E. CHARRO,[2] A. ZACARÍAS,[2] M. CASTRO[2]**

[1] *Northwestern University Medical School, Department of Molecular Pharmacology and Biological Chemistry, 303 East Chicago Avenue, Chicago, Illinois 60611-3008*
[2] *Departamento de Física y Química Teórica, UNAM, D.F. 04510, México*

Received 13 May 1998; accepted 10 November 1998

ABSTRACT: The vibrational spectra of the nucleic acid bases adenine, thymine, guanine, and cytosine are calculated in the frame of density functional theory (DFT). In particular we use the Kohn–Sham scheme with gradient corrections for exchange and correlation to determine normal modes, frequencies, and intensities. The DFT results are found to be in good agreement with the experiment. Our computations provide assignments for IR, Raman, and neutron inelastic scattering spectroscopies; yield characteristic vibrational fingerprints of each compound for its identification in larger systems; and show general vibrational trends of nucleic acids. The Kohn–Sham scheme is further applied to obtain the spectra of the Watson–Crick pairs adenine-thymine and guanine-cytosine. A large number of monomeric vibrations are recognized in dimers; characteristic vibrations of pairs, which are mainly attributed to hydrogen bridges, are quantified according to changes in normal modes and frequency shifts. Binding and zero-point vibrational energies are analyzed to establish the stability of the complexes and discuss the quality of the energetic calculations.
 J Comput Chem 20: 511–530, 1999

Keywords: vibrational spectrum; nucleic acids; hydrogen bridges; density functional theory; gradient corrections

* Permanent address: Instituto de Física, UNAM, Apdo. Post. 20-364, D.F. 01000, México

** Permanent address: Dept. de Química-Física, Universidad de Valladolid, 47005 Valladolid, Spain

Correspondence to: R. Santamaria; e-mail: rso@fenix.ifisicacu.unam.mx

 CCC 0192-8651 / 99 / 050511-20

Introduction

The identification of a variety of compounds by vibrational spectroscopy has long been recognized.[1] Nevertheless, for large molecules the interpretation of the resulting spectra becomes difficult and frequently uncertain. Major problems to be solved are, for instance, the minimization of intermolecular interactions between the compound and its environment, which may either be an inert gas matrix, a solution, or a crystalline solid. Also, the presence of impurities (rotamers, tautomeric conformers, or associated molecules) and anharmonic effects, as well as improper force fields of model compounds taken as first approximations at a given temperature, represent some of the important factors that complicate experimentation.

A number of experiments were performed for DNA fragments to determine their vibrational modes and wave numbers.[2–8] Most of them were based on optical techniques such as IR absorption or Raman scattering and fewer on methods like neutron inelastic scattering (NIS). However, noticeable discrepancies still exist. It is our purpose in this work to elucidate the harmonic vibrational spectra of the nucleic acids and their pairs. It is important to remark that new techniques in many-body theory offer alternative tools for chemical research.[9] Thus, accurate structures, energies, and other electronic features can be predicted not only for small but also for medium sized molecules. Furthermore, the nucleic acid bases (NABs) constitute the most important pieces of genetic information. Hence, the study of their vibrational spectra becomes of interest, especially in the gas phase where they are free of external forces. In fact, a vibrational analysis of adenine (A), guanine (G), thymine (T), and cytosine (C) provides a departure point for comparison with other compounds of similar nature or their analogues in a different environment. Thus, it is common to find studies that involve protonated species,[10,11] tautomeric forms,[12–20] or nucleic bases embedded in crystal structures[2–5,8] or in solution.[21–23]

For further insight we focus attention on those changes produced by hydrogen-bond formation. The quantification of the effects of hydrogen (H) bridges in the spectra of NABs is of great interest for the spectroscopic identification of nucleic acid pairs inside the double DNA helix. This is the reason we also include in our analysis the adenine-thymine (AT) and guanine-cytosine (GC) Watson–Crick complexes.

Method

Most quantum mechanical schemes that yield accurate results are complicated. Hence, a compromise between accuracy and computational efficiency is necessary for the proper study of medium sized and large compounds.

The simplest *ab initio* method is the Hartree–Fock (HF) approximation.[24] It is capable of giving the gas phase vibrational spectra of the nucleic acids with relatively good accuracy after introducing one or more scale factors fitted to experimental data from model compounds to account for correlation effects.[25] However, scale factors may not be independent of the environment, leading to different results for a single compound. Also, the HF method requires large basis sets for satisfactory results and it becomes quite a time-consuming approach. An alternative scheme is density functional theory (DFT),[26,27] which already involves correlation effects. Such a theory has been applied to a variety of chemical problems.[28,29] The computational effort in DFT calculations formally scales as N^4, where N is the number of electrons, due to the evaluation of four-center integrals as in HF theory. However, the DFT approach becomes computationally more expensive because many integrations have to be done numerically. In practice the exponent gets smaller than 3 for both methods with a resort to cutoffs. Such a scaling favors the use of DFT (which involves correlation) to study the vibrational spectra of the DNA nucleic acids.

We apply the Kohn–Sham version of DFT.[27] In this formulation the wave function and electron density play a major role. The exact wave function in the Kohn–Sham reference system of noninteracting particles is represented by a single determinant, and exchange-correlation effects are calculated with the use of electron density functionals in contrast to traditional quantum mechanical methods where correlation is computed with an elaborate multideterminant wave function or evaluated in a troublesome perturbative process.[30,31]

The accuracy in DFT is determined to a great extent by the particular functional used. Therefore, we consider the simple and well-known local spin density (LSD) approximation, which is based on the uniform electron gas approach.[32] Because the LSD approximation overestimates the strength of

H bonds, it requires amendments. Then density gradient corrections of the exchange-correlation term due to Becke (B) and Perdew (P)[33] are additionally incorporated into the LSD scheme to correctly account for electron inhomogeneities and energetics.[34] Calculations are performed in a self-consistent field way as implemented in the computational package Gaussian 94.[35] Such a package only needs orbital basis sets for specifying the orbital functions themselves and the electron density. We use the 6−311G orbital basis sets, which are of sufficient quality to investigate the vibrational spectra of compounds considered in this work. Take note that for the case of nucleic acid monomers it is possible to use larger basis sets,[36] such as those that include polarization functions; but for the case of dimers eventual problems are expected to arise due to computational limitations. Under such circumstances pair complexes would demand smaller basis sets. Then the use of sets like 6−311G permits a reasonable balance between accuracy on the one hand and the analysis of monomers and dimers within the same computational framework on the other. In the Gaussian 94 package the level of theory used here is stated as BP86/6−311G.

Fully relaxed ground state structures are determined with the computation of analytic gradients, and the Berny optimization process[37] was used as the search algorithm in terms of redundant internal coordinates. The threshold for gradient convergence is set to 0.00035 hartrees per bohr (hartree/b) for each Cartesian component of every atom, and computations are carried out on an Origin 2000 computer.

We note that a similar DFT approach was applied to investigate the structures and energetics of DNA fragments and encouraging results were achieved.[38,39] It is also important to mention that studies exist that are related to the vibrational spectra of organic compounds where the LSD scheme (without the need of gradient corrections) performs well. For instance, after a systematic study of main-group molecules,[40,41] it was shown that vibrational frequencies are relatively insensitive to the DFT level of theory and basis set used. However, when dealing with either H-bonded complexes,[42] transition metal monoligands,[43] or s-radicals,[44] gradient corrections become important. Therefore, the gradient corrected approximation employed in this work should for the most part be of sufficient accuracy to assist us in the vibrational assignment task but with critical analysis on these vibrations where H bridges are involved.

Vibrational Spectra of NABs

ADENINE

The adenine molecule (Fig. 1) has 39 normal vibrations[45] with 27 in the plane (Table I). All frequencies are real and indicate a true ground state. Given the nature of the spectrum, we only discuss what we consider to be the most important assignments.

Experimental results are also reported in Table I. Quantification of deviations between theory and the experiment is not easy because even when frequencies from different experiments are similar, assignments may be different or ambiguous. Nevertheless, experimental results may complement each other because, for example, IR and UV resonance Raman scattering are known to give accurate high wave numbers (approximately above

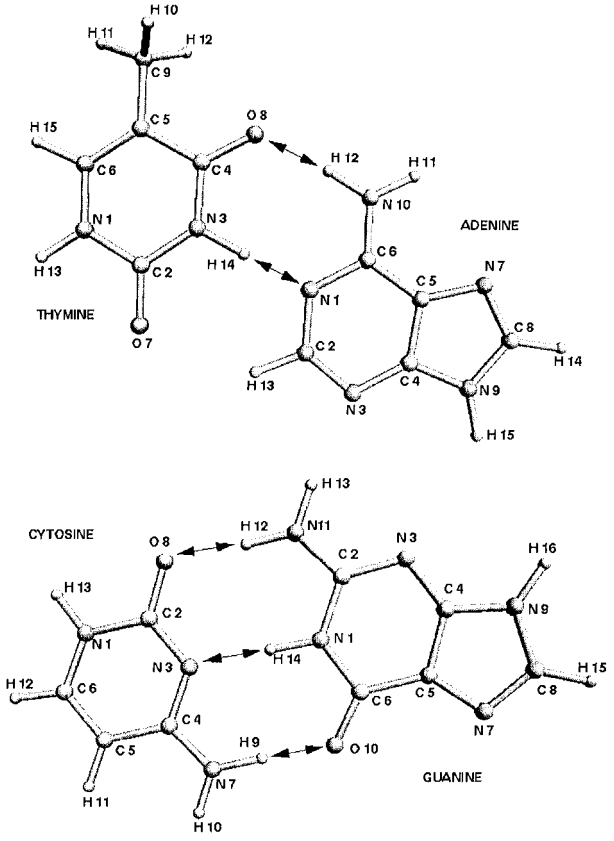

FIGURE 1. The nucleic acid bases adenine, thymine, guanine, and cytosine. Hydrogen bridges are indicated with double-headed arrows.

500 cm^{-1}) for in-plane vibrations of heavy atoms, while NIS gives accurate low wave numbers for out of plane vibrations of hydrogen.[4] Thus, care should be taken when comparing theoretical results with experiments in different intervals.

In the very low frequency region, which is below 200 cm^{-1}, only NIS experimental values are reported because of the existence of lattice vibration noise in the IR and Raman experiments.[3] We found a single DFT mode due to the whole torsion

TABLE I.
Vibrational Spectrum of Adenine.

		Experiment					
No.	BP	Ref. 2[a]	Ref. 3[b]	Ref. 4[c]	Ref. 5[d]	Plane	Assignment
1	161 (9.2)			184		Out	Tors molecule, wag C6 — NH_2
2	209 (7.3)	240	249	194 / 238		Out	Butt molecule
3	261 (14.2)		327		240	In	Bend molecule, bend C6 — NH_2
4	296 (1.0)	316	337	310 / 331		Out	Tors molecule
5	499 (4.8)		530			In	Bend molecule (sqz group N3 — C4 — N9), bend C6 — NH_2
6	511 (227.3)	518				Out	Wag N10 — H11, N10 — H12
7	514 (3.2)		543	529	535	In	Def R6 (S sqz group N1 — C6 — C5, C2 — N3 — C4)
8	553 (1.1)	560		550		Out	Rock NH_2, wag C2 — H, C8 — H
9	559 (6.0)	594		624		Out	Rock NH_2, wag C2 — H
10	600 (1.6)	615	622	620	620	In	Def R5, R6
11	640 (202.8)					Out	S wag N9 — H, C8 — H
12	649 (12.6)					Out	A wag N9 — H, C8 — H
13	674 (1.0)	686	660 / 684	655 / 686		Out	Def R5, R6, A wag N9 — H, N10 — H12
14	693 (3.4)		724	721	722	In	Breath molecule
15	782 (31.3)					Out	Def R5, R6 (wag group C4 — C5 — C6), wag C8 — H
16	807 (0.8)		797	850		Out	Def R5, R6 (wag group C4 — C5 — C6), wag C8 — H
17	853 (8.1)	861			848	In	Def R6 (sqz group N1 — C2 — N3)
18	914 (13.0)		951	942	898	In	Def R5 (sqz group N7 — C8 — N9)
19	917 (9.2)	882 / 935	849	943		Out	Wag C2 — H
20	969 (1.9)	963	913		940	In	Def R6 (str N1 — C6), S bend N10 — H11, N10 — H12
21	1053 (26.7)	1053 / 1027		1032	1023	In	Def R5 (str C8 — N9), A bend C8 — H, N9 — H
22	1103 (6.8)	1127	1126	1129	1126	In	Def R5, R6, bend C2 — H, C8 — H, N9 — H
23	1202 (8.3)			1168	1164	In	Def R5, R6 (str C5 — N7), S bend N10 — H11, N10 — H12
24	1213 (40.8)	1227	1157 / 1234	1234 / 1248	1235	In	Def R5, R6, S bend C8 — H, N9 — H
25	1260 (89.3)	1253	1253		1248	In	Def R5, R6 (str C2 — N3, C5 — N7), A bend C8 — H, N9 — H
26	1292 (48.9)	1282			1307	In	Def R5, R6 (str N1 — C2, C5 — N7), bend C8 — H
27	1328 (14.5)	1320	1309	1329	1331	In	Def R5, R6, bend C2 — H, C8 — H, N9 — H, N10 — H11
28	1355 (14.6)	1380	1335	1366	1370	In	Def R5, R6 (str C4 — N9, N7 — C8), bend all H
29	1387 (0.4)		1371 / 1451			In	A bend C2 — H, N9 — H
30	1421 (36.9)	1425	1421		1418	In	Def R5, R6, S bend all H
31	1450 (11.1)	1488 / 1476	1482	1483 / 1471	1462 / 1482	In	Def R5, R6, bend C2 — H, C8 — H, str C6 — NH_2
32	1552 (93.7)	1590 / 1615	1508	1598	1510 / 1597	In	Def R5, R6 (str N3 — C4), bend N9 — H, N10 — H11

(Continued)

TABLE I.
(Continued)

No.	BP	Experiment				Plane	Assignment
		Ref. 2[a]	Ref. 3[b]	Ref. 4[c]	Ref. 5[d]		
33	1583 (69.3)	1630	1604 / 1638		1612	In	Def R6 (str C5 — C6), sciss NH_2
34	1634 (456.8)	1650	1673	1680	1675	In	Sciss NH_2, str C6 — NH_2
35	3130 (24.9)	3026			3038	In	Str C2 — H
36	3194 (1.3)	3126			3125	In	Str C8 — H
37	3509 (78.9)	3442				In	S str N10 — H11, N10 — H12
38	3568 (51.4)	3490				In	Str N9 — H
39	3655 (38.7)	3558				In	A str N10 — H11, N10 — H12

Bend, bending; breath, breathing; butt, butterfly; def, deformation; rock, rocking; sciss, scissoring; sqz, squeezing; str, stretching; tors, torsion; umb, umbrella; wag, wagging; A, asymmetric; S, symmetric; R5, five-membered ring; R6, six-membered ring. BP indicates the exchange-correlation functional (level of density functional theory) used in this work; refer to the text for details.
Frequency units in cm^{-1} and intensities, which appear in parentheses, in km / mol. Bonds inside parentheses indicate the main source of deformation for the given vibrational mode.
[a] IR results of adenine in an inert gas matrix.
[b] IR and Raman combined results of adenine in crystalline powder.
[c] Neutron inelastic scattering and UV resonance Raman scattering combined results of adenine in polycrystalline samples and solvent.
[d] IR and Raman frequencies (without assignments given in the original reference) of polycrystalline adenine.

of the molecule together with a wagging of the NH_2 group.

Out of plane torsions occur in the low frequency range, which is below 950 cm^{-1}. They are characterized by coupled movements of all atoms in the molecule. Many torsions also show wagging of hydrogen bonds. In a similar fashion, in-plane modes describe mixed collective motion resulting in complex deformation of six- and five-membered rings. In this last case, the main source of deformation is attributed to angle squeezing.

Among the first modes that call attention in the low frequency range is the butterfly-like motion. In this mode the six- and five-membered units simulate wings. According to our calculation it has a frequency of 209 cm^{-1} compared to experimental frequencies that appear in the interval 194–249 cm^{-1}. Another interesting mode, this time of in-plane nature, is mode 14. It shows the expected breathing of adenine. Experiments yield values of 721 and 724 cm^{-1}, and our result differs by 4%.

Note in Table I the existence of assignments that describe dominant wagging oscillations of hydrogen bonds. They classify as out of plane motions.[45] Modes 11 and 12 (with no experimental values associated) report symmetric and asymmetric waggings of C8—H and N9—H. Mode 19 only shows C—H wagging.

In the intermediate region, which is above 950 cm^{-1} and below 1700 cm^{-1}, the spectrum is characterized by in-plane deformations of mostly five- and six-membered rings. This region is where covalent stretches start to appear and the bendings of H bonds are abundant. In the last part of this region we find scissoring of NH_2. Scissoring frequencies compare well with the experiment, giving deviations lower than 3%.

In the interval 3000–3700 cm^{-1} vibrations 35–39 arise exclusively from hydrogen stretches. Symmetric and asymmetric motions are predicted for NH_2. Observe that other stretching vibrations appeared before reaching the tail of the spectra (those responsible for in-plane ring deformations at the intermediate region); however, they involved no H stretch. In general, DFT assignments and frequencies in the tail region agree well with IR measurements from the inert gas matrix[2] because our differences are of the order of 3%. At this high frequency interval other theoretical schemes[12] give similar assignments to ours but larger frequencies, and semiempirical computations yield good frequencies but incorrect assignments.

If a comparative analysis is extended to the whole spectrum of adenine, DFT results present an average deviation of 3% with respect to the IR values of ref. 2. The largest single deviation (13%) occurs for mode 2 in the low frequency regime.

There are vibrations along the spectrum where the amino group NH_2 clearly participates, for example, in rocking (modes 8 and 9), bending (modes 20 and 23), and scissoring movements (modes 33

and 34). Similar contributions are expected for the NH_2 groups of cytosine and guanine.

As an illustration of the utility of our spectrum, in the last column of Table I we present IR and Raman measurements from Majoube.[5] Because no assignment was done by Majoube in that study, we simply matched his frequencies and ours (see also ref. 4). We obtained an average deviation of 2.3%. As a result, a combined theoretical–experimental spectrum is shown that agrees well with those from other authors who have recourse to different techniques for establishing assignments.[2–4]

Finally, the modes with largest intensities are in decreasing order 34, 6, 11, 32, and 25. The simplest vibrations are the first two, which are mainly related to hydrogen oscillations in the amino group.

THYMINE

This molecule (Fig. 1) consists of 39 vibrations (Table II). Comparative frequencies for in-plane oscillations obtained from *ab initio* (MINI-1) and semiempirical (PM3)[12] computations are included in Table II. The agreement between these last two schemes is not good at high wave numbers. Therefore, the comparison in that spectral region should be performed with care.

The mixed motion of ring atoms leads to a great number of deformations. Out of plane deformations extend as far as 1058 cm^{-1}, contrary to adenine where they extend to 917 cm^{-1}.

At very low frequencies three vibrations show values smaller than 200 cm^{-1}. They yield molecular torsions. The first two modes also present CH_3 rocking, and the third one simulates a kind of butterfly with the two C—O bonds simulating wings. Experiments in the solid state may face assignment problems in this interval that are due to the lattice phonons.

In the low frequency region from 200 cm^{-1} to 1058 cm^{-1}, where we have the last out of plane oscillation, ring deformations are observed with wagging of hydrogens. The simplest modes (4, 6, and 10–12) involve either bending of C—O or C—CH_3 or wagging of N—H bonds. Mode 13 is related to the breathing of the molecule. In A the breathing was found at 693 cm^{-1} and experimentally assigned at 721–724 cm^{-1}, while in T it was predicted at 718 cm^{-1}. Thus, we can to a certain extent assume that our DFT value is relatively correct. The remaining in-plane and out of plane vibrations classify as ring deformations.

In the intermediate wave number range from 1100 to 1700 cm^{-1} we again observe many deformations. All are in the plane. The first important C—C, C—N, and C—O stretches appear in this interval together with many H bends. The simplest modes to visualize are 27, 29, and 30, which show umbrella-like motion and scissoring of CH_3. In fact, the presence of CH_3 along the spectrum is well evidenced through modes showing rocking, bending, and wagging of its hydrogens and stretching of the C5—CH_3 bond.

The tail of the spectral region from 2900 cm^{-1} onward is only characterized by H stretches. The previous good agreement (3%) between our values and experimental measurements in A gives us confidence to assume that results for the same kind of vibrations in T are reliable to a similar extent. Thus, in this interval the MINI-1 method overestimates frequencies and some PM3 assignments appear inverted. If we consider the entire spectrum, PM3 and MINI-1 give average deviations of 5.6 and 7.2% with respect to our values. However, single deviations of PM3 may be as large as 21%, contrary to those of MINI-1 that do not exceed 15%.

Finally, the vibrations with the largest intensities correspond to modes 15 and 31–33. In some of them C—O stretches are a common factor. Such modes, together with these of CH_3, may be considered good markers of thymine to discriminate from the other NABs.

GUANINE

Guanine (Fig. 1) has 42 normal modes. Table III reports our values and assignments and experimental data from ref. 6. Guanine presents experimental difficulties in establishing its vibrational spectrum. The main reason is the formation of numerous hydrogen bonds with its surrounding that lead to distorted vibrations.[7]

From the structural point of view, guanine resembles adenine and some similar vibrations are expected to occur.

Below 200 cm^{-1} we found three different torsions of the molecule. The first two additionally include wagging of either C2—NH_2 or C6—O bonds. The third mode at 194 cm^{-1} represents a deformed butterfly motion due to the C6—O bond that basically remains static. The IR experiment yielded a value of 204 cm^{-1} for this last mode and in adenine it was at 209 cm^{-1}.

TABLE II.
Vibrational Spectrum of Thymine.

No.	BP	Theory[a] MINI-1	Theory[a] PM3	Plane	Assignment
1	110 (0.0)			Out	Tors ring, rock CH_3
2	125 (0.6)			Out	Tors ring, rock CH_3
3	167 (2.3)			Out	Tors ring (wag group C2 — N3 — C4)
4	273 (3.3)	233	304	In	Bend C5 — CH_3
5	306 (3.3)			Out	Tors ring, wag C5 — CH_3
6	365 (17.4)	343	346	In	A bend C2 — O, C4 — O
7	407 (9.2)			Out	Tors ring, wag C6 — H
8	452 (17.0)	414	429	In	Sqz ring
9	526 (7.1)	504	499	In	Def ring (sqz group C2 — N3 — C4), A bend N1 — H, C2 — O
10	583 (1.1)	546	554	In	S bend C2 — O, C4 — O, C5 — CH_3
11	628 (81.6)			Out	Wag N1 — H
12	696 (2.0)			Out	Wag N3 — H
13	718 (3.2)	703	822	In	Breath molecule
14	725 (39.0)			Out	Def ring (wag group N1 — C2 — N3), S wag N1 — H, N3 — H
15	779 (140.8)			Out	Def ring, wag N3 — H
16	780 (4.3)	772	749	In	Def ring (sqz group C6 — N1 — C2), str C5 — CH_3
17	879 (15.5)			Out	Def ring, wag C6 — H
18	933 (8.2)	937	963	In	Def ring, A bend N3 — H, C6 — H, S bend C9 — H10, C9 — H11, C9 — H12
19	1007 (1.4)			Out	Def ring, S wag C9 — H10, C9 — H11, C9 — H12
20	1058 (3.5)			Out	Def ring, wag C9 — H10, C9 — H11, C9 — H12
21	1130 (10.6)		1207	In	Def ring (str N3 — C4)
22	1172 (87.3)	1145	1132	In	Def ring, A bend N1 — H, C6 — H
23	1208 (37.4)	1287	1267	In	Def ring (str N1 — C6), str C5 — CH_3
24	1333 (35.4)	1224	1362	In	S bend N3 — H, C6 — H
25	1365 (19.0)	1379	1362	In	A bend N3 — H, C6 — H
26	1381 (73.0)	1431	1380	In	Def ring (str C4 — C5), S bend N1 — H, N3 — H
27	1413 (8.6)	1456	1402	In	Umb CH_3
28	1446 (13.4)		1509	In	Def ring (str N1 — C6, C4 — C5), bend N1 — H
29	1464 (11.1)	1529	1436	In	Sciss CH_3
30	1483 (5.2)	1550		In	Sciss C9 — H10, C9 — H12
31	1609 (124.7)	1793	1941	In	Def ring, str C4 — O, A bend N3 — H, C6 — H
32	1646 (218.0)	1854	1816	In	Def ring (str C4 — C5, C5 — C6), S bend N1 — H, C6 — H
33	1680 (654.1)	1819	1898	In	Def ring, str C2 — O, A bend N1 — H, N3 — H
34	2949 (25.2)	3306	3179	In	S str C9 — H11, C9 — H10, C9 — H12
35	3010 (13.9)	3459	3087	In	A str C9 — H10, C9 — H12
36	3036 (17.4)	3459	3087	In	A str C9 — H11, C9 — H10, C9 — H12
37	3121 (10.3)	3402	3021	In	Str C6 — H
38	3502 (39.9)	3868	3378	In	Str N3 — H
39	3543 (66.1)	3898	3409	In	Str N1 — H

See footnotes of Table I for abbreviations. Note that the CH_3 radical owns two equivalent hydrogens, H10 and H12, in Figure 1. They are located above and under the molecular plane in a symmetric fashion. The third hydrogen, H11, is found in the molecular plane. In this way, it is possible to speak about symmetric and asymmetric vibrational modes that involve such hydrogens as for example modes 34–36.

[a] *Ab initio* (MINI-1) and semiempirical (PM3) results for in-plane vibrations from ref. 12.

The low frequency region, with the last out of plane vibration taken as the upper limit, ends at 782 cm^{-1}. That region for A and T extends as far as 917 and 1058 cm^{-1}, respectively. We find out-of-plane collective motions that lead to ring deformations and wagging of H bonds. The simplest modes are 4, 9, and 15; with the wagging of N—H bonds only. Their frequencies differ approximately 9% from the experiment, and they show the largest intensities in the interval. Modes 5 and 6, which involve bending of C—NH_2 and C—O bonds, differ substantially from the experiment and show

TABLE III.
Vibrational Spectrum of Guanine.

No.	BP	Experiment Ref. 6[a]	Plane	Assignment
1	134 (3.7)		Out	Tors molecule, wag C2 — NH_2
2	169 (0.4)		Out	Tors molecule, wag C6 — O
3	194 (2.9)	204	Out	Def butt molecule
4	264 (89.2)	243	Out	Wag N11 — H12
5	306 (3.0)	340	In	Bend molecule, S bend C2 — NH_2, C6 — O
6	319 (2.4)	387	In	A bend C2 — NH_2, C6 — O
7	354 (13.3)	356	Out	Def R5, R6 (wag group C2 — N3 — C4)
8	476 (1.0)	493	In	Def R6 (S sqz group C2 — N3 — C4, N1 — C6 — C5)
9	501 (195.3)	501 / 578	Out	Wag N11 — H13
10	507 (0.6)	581	In	Def R6 (sqz group C2 — N1 — C6), bend N11 — H13
11	600 (8.6)	639	In	Breath molecule
12	615 (3.4)	547	Out	Def R5, R6, A wag N1 — H, N9 — H
13	639 (16.4)	740	In	Def R6 (sqz group N1 — C2 — N11), bend C6 — O
14	640 (0.3)	654	Out	Def R5, R6 (wag group N7 — C8 — N9), A wag C8 — H, N9 — H
15	646 (175.3)	690	Out	S wag N1 — H, N9 — H
16	682 (25.4)	703	Out	Def R5, R6 (wag group N1 — C2 — N3)
17	707 (66.2)	790	Out	Def R5, R6, wag N1 — H
18	756 (66.7)	778	Out	Def R5, R6 (wag group N1 — C6 — C5), wag C8 — H
19	782 (6.6)	838	Out	Def R5, R6 (wag group C4 — C5 — C6), wag C8 — H
20	795 (6.2)	874	In	Def R5, R6 (sqz group C4 — N9 — C8), bend N11 — H13
21	915 (8.8)	924	In	Def R5 (sqz group N7 — C8 — N9)
22	993 (2.5)	1030	In	Def R6 (str N1 — C2), S bend N1 — H, N11 — H12, N11 — H13
23	1013 (29.5)	1112	In	Def R6 (str N1 — C6), S bend N11 — H12, N11 — H13
24	1038 (29.8)	1149	In	Def R5 (str C8 — N9), A bend C8 — H, N9 — H
25	1080 (22.4)	1136	In	Def R6 (str N1 — C6), S bend N11 — H12, N11 — H13
26	1130 (16.0)	1191	In	Def R5, R6 (str C4 — N9), S bend C8 — H, N9 — H
27	1253 (0.9)	1275	In	Def R5 (str N7 — C8), bend C8 — H
28	1296 (41.1)		In	Def R5, R6 (S str C4 — N9, C5 — N7), bend N1 — H
29	1323 (32.4)	1380	In	Bend N1 — H, N9 — H, N11 — H13
30	1333 (24.9)	1430	In	Def R5 (str C8 — N9), A bend C8 — H, N9 — H
31	1376 (21.1)	1356	In	Def R5, R6 (A str N1 — C2, C4 — C5), S bend N9 — H, N11 — H13
32	1428 (0.2)	1487	In	Def R5, R6 (str N7 — C8), bend all H
33	1483 (39.6)	1505	In	Def R5, R6 (S str N1 — C2, C4 — C5), bend N11 — H13
34	1541 (109.4)	1565	In	Def R5, R6 (str N3 — C4, C4 — N9), bend N9 — H
35	1570 (442.0)	1637	In	Def R6 (str C2 — N3), bend N1 — H, sciss NH_2
36	1638 (370.7)	1609	In	Sciss NH_2
37	1686 (443.9)	1689 / 1702	In	Def R6 (str C5 — C6), str C6 — O, bend N1 — H, sciss NH_2
38	3196 (1.5)	3116	In	Str C8 — H
39	3480 (15.5)	2717	In	Str N1 — H
40	3530 (90.8)	3202	In	S str N11 — H12, N11 — H13
41	3563 (40.3)	2926	In	Str N9 — H
42	3670 (35.9)	3306	In	A str N11 — H12, N11 — H13

See footnotes of Table I for abbreviations.
[a] IR and Raman combined results of polycrystalline guanine.

very small intensities. Compared with the other NABs, the breathing of G at 600 cm^{-1} appears with the lowest frequency: A had a value of 693 cm^{-1}, T of 718 cm^{-1}, and C of 731 cm^{-1}. For the breathing case, experimental and theoretical values are close to each other; and the larger the NABs, the lower the breathing frequency.

In-plane ring deformations characterize the intermediate frequency range from 795 to 1686 cm^{-1}. Most of them are attributed to N—C and C—C

stretches, which is in contrast to the squeezing or wagging of angles that is the main source of deformation in the previous region. The bending of H bonds is also abundant here. Mode 29 is one of the simplest vibrations with only N–H bending. The scissoring motion is present in modes 35–37, which is occasionally mixed with deformations of the six-membered ring. They show the largest intensities in the whole spectrum. The DFT results are in good agreement with the experiment at this interval (except modes 20 and 24 with deviations of 10%).

The NH_2 group participates actively in G as it does in A or CH_3 in T. The modes of NH_2 along the spectrum (see, e.g., modes 1, 4–6, 9, 23, 25, and 35–37) represent the fingerprints of the amino group in the G compound.

At the tail of the spectrum one finds H-bond stretching. Hydrogens of NH_2 take part in symmetric and asymmetric forms. In this region differences exist between the experiment and theory. From our experience on A, we infer that natural bonds of G changed in the polycrystalline phase due to the environment (see also ref. 7). The influence of the environment on G is expected to have also changed frequencies in previous intervals. Thus, the above comparisons between theory and experiment should not be used to establish the accuracy of either result but should rather used to show expected trends of vibration of guanine in different conditions.

Finally, different vibrations and frequencies appear when tables of the G and A purines are compared. The different vibrations are important to discriminate between the two species. As an example, we have those vibrations of G that involve N1—H or C6—O bonds not present in A, or the C2—H bond of A not present in G.

CYTOSINE

Cytosine (Fig. 1) has 33 normal vibrations; 23 of them are in the molecular plane (Table IV). The *ab initio* results and experimental measurements from Ar and N_2 matrix-isolation spectroscopies, as well as from cytosine in the crystalline solid, are included in Table IV. Ar and N_2 matrices are transparent to IR radiation and, in contrast to cytosine in the solid, they minimize interaction with the pyrimidine compound.[8] The *ab initio* results of ref. 8 were used to interpret and provide assignment for the experimental spectrum. Such computations assumed a planar structure and the scaling of frequencies was necessary to bring theoretical frequencies into close agreement with the experiment, especially for N—H and O—H stretches. However, other kinds of vibrations are expected to be affected due to the use of a single (0.91) scale factor.

Cytosine only presents two vibrational modes in the very low frequency region. They show molecular torsion with wagging of C_2—O, but the second one also includes the wagging of C4—NH_2.

The next region from 300 to 928 cm^{-1} shows several molecular deformations. In-plane ring deformations are attributed to angle squeezings while out of plane ones are associated with angle waggings. The interesting vibrations in the region are the symmetric and asymmetric bending of C2—O with C4—NH_2 (modes 3 and 6) and the isolated wagging of N—H and C—H bonds (modes 9, 10, and 16). Most of them compare favorably with values obtained from Ar and N_2 experiments and HF computations but differ from the results of cytosine in the solid. The breathing of the molecule at 731 cm^{-1} lies in the experimental range of 716–780 cm^{-1} and compares well with the HF result of 743 cm^{-1}.

In the intermediate frequency region ring deformations are still present. Instead of the previous wagging and squeezing movements of angles, C—N and C—C stretches are the current sources of in-plane deformation. The stretching motion of the C—O bond is the last one to show. Also, the bending of H bonds appears in large numbers. Modes 20–21 and 24–25 involve bending of practically all hydrogens. The simplest vibrations, modes 18–19 and 21–22, are coupled C—H and N—H bending motions. A general good agreement exists among our results and the values from the different methodologies presented in Table IV.

Stretching motions are predicted at the tail of the spectrum from 3100 to 3700 cm^{-1}. The symmetric vibration of amino hydrogens appears with lower frequency than its asymmetric counterpart. In a similar fashion as for G in its crystal form, experiments of C in the crystal state show differences with DFT values and with matrix-isolation measurements. Compare, for example, modes 32–33 of C with 41–42 of G where frequencies are similar for similar vibrations. Thus, we reach the same conclusions for C in the solid as previously reached for polycrystalline guanine.

Finally, the amino group takes part in several important vibrations along the spectrum of cytosine. With regard to the intensities of the vibrations, modes 27–28, 24, and 5 give the largest values. Modes that involve scissoring of amino

hydrogens or bond stretches, frequently of the C–O type, yield the largest intensities.

Comparative Analysis of Vibrational Spectra

It becomes of interest to make a spectral comparison in order to find similar vibrations among the different compounds. We consider that normal modes become similar when their frequencies are almost identical within a maximum deviation of 20 cm^{-1} and exhibit the same kind of vibration. The identification of similar vibrational patterns in species that are in principle different is a difficult task. Nevertheless, common elements in the nucleic acids like five- and six-membered rings or functional groups usually show vibrational similarities. The importance of identifying similar modes is because they assist us in discriminating between different families of compounds, in this context purines from pyrimidines. In Table V we report vibrations that are similar in the purine molecules (column A = G) and in the pyrimidine molecules (column T = C). Not surprisingly, purines show a larger number of entries than

TABLE IV.
Vibrational Spectrum of Cytosine.

No.	BP	Experiment[a] Ar	N_2	KBr	Theory[b] HF	Plane	Assignment
1	139 (3.2)	197			167	Out	Tors molecule, wag C2 — O
2	197 (9.7)	232			218	Out	Tors molecule, S wag C2 — O, C4 — NH_2
3	342 (4.2)		360	425	354	In	A bend C2 — O, C4 — NH_2
4	409 (19.7)		409	440	442	Out	Tors molecule, A wag C6 — H, N7 — H10
5	415 (148.1)	330	330	500	452	Out	Tors molecule, S wag C6 — H, N7 — H10
6	503 (3.0)		532	545	525	In	S bend C2 — O, C4 — NH_2
7	530 (2.0)	537	537	530	535	In	Def ring (sqz group N3 — C4 — C5)
8	560 (1.8)	575		600	582	In	Def ring (sqz group C2 — N3 — C4)
9	576 (89.3)	614		700	598	Out	Wag N7 — H9
10	667 (79.3)	637	643	760	692	Out	Wag N1 — H
11	712 (21.9)	623			765	Out	Def ring (wag group N3 — C4 — C5), S wag N1 — H, C5 — H
12	731 (3.6)	747	716	780	743	In	Breath molecule
13	739 (97.9)	818	726	820	861	Out	Def ring (wag group N1 — C2 — N3), wag N1 — H
14	764 (38.0)	781		791	882	Out	Def ring (wag group N3 — C4 — C5), S wag C5 — H, C6 — H
15	884 (3.9)			885	891	In	Def ring (sqz group N3 — C4 — N7), S bend all H
16	928 (7.1)	1087	1110	996	1078	Out	A wag C5 — H, C6 — H
17	948 (2.9)			1010	972	In	Def ring (sqz group N1 — C6 — C5)
18	1035 (53.3)	1090		1150	1072	In	S bend N7 — H9, N7 — H10
19	1103 (2.4)	1124	1134	1100	1107	In	A bend N1 — H, C5 — H
20	1192 (20.1)	1192	1202	1238	1205	In	Def ring (str N1 — C2, C2 — N3), bend all H
21	1215 (52.8)	1244	1245	1278	1236	In	Bend all H
22	1339 (28.1)	1337	1341	1363	1354	In	S bend N1 — H, C5 — H, C6 — H
23	1396 (36.4)	1422	1425	1463	1420	In	Def ring (str N1 — C6, N3 — C4), bend all H
24	1473 (170.1)	1475	1479	1505	1482	In	Sciss NH_2, str C4 — NH_2, S bend all H of ring
25	1495 (72.8)	1539	1541	1540	1544	In	Def ring (str C4 — C5), bend all H
26	1613 (12.6)	1595	1601	1700	1642	In	Def ring (str C5 — C6), sciss NH_2
27	1631 (643.5)	1656	1657	1640	1663	In	Def ring (str C5 — C6), sciss NH_2
28	1660 (390.4)	1717	1712	1662	1773	In	Def ring, str C2 — O, sciss NH_2, bend N1 — H
29	3116 (4.6)				3086	In	A str C5 — H, C6 — H
30	3142 (11.5)				3123	In	S str C5 — H, C6 — H
31	3505 (69.2)	3441	3439		3442	In	S str N7 — H9, N7 — H10
32	3517 (32.5)	3471	3460	2850	3466	In	Str N1 — H
33	3654 (28.9)	3565	3562	3380	3562	In	A str N7 — H9, N7 — H10

See footnotes of Table I for abbreviations.
[a] Experimental results from ref. 8 of cytosine in Ar and N_2 matrices and in the crystalline solid.
[b] Scaled Hartree–Fock (HF) results from ref. 8, assuming a planar structure of cytosine.

TABLE V.
Similar DFT (BP) Vibrational Modes among Different Nucleic Acid Bases and Dimers.

A = G			C = T			A = AT			T = AT		G = GC		C = GC		
1	2	[184]	(12)	(13)	[731]	1	9	[184]	1	5	1	5	2	10	[232]
(2)	(3)	[230]	(28)	(33)	[1714]	(2)	(11)	[230]	2	8	(2)	(9)	(8)	(23)	[575]
6	9		(32)	(38)	[3465]	4	13	[323]	(4)	(12)	3	10	(11)	(31)	[623]
11	15					(5)	(19)		5	14	7	13	(13)	(32)	[772]
12	14					(7)	(20)	[536]	(7)	(17)	(8)	(17)	(14)	(36)	[781]
15	19					12	28		(8)	(18)	(10)	(18)	16	40	[1098]
18	21	[946]				14	30	[722]	(9)	(21)	14	26	21	52	[1244]
(21)	(24)	[1037]				15	34		(10)	(24)	(16)	(30)	23	59	[1423]
27	30	[1319]				16	36	[823]	11	26	(18)	(33)	(25)	(63)	[1540]
30	32	[1423]				18	40	[946]	13	32	19	35	(29)	(74)	
(32)	(34)	[1578]				21	45	[1037]	16	35	20	37	30	75	
(33)	(35)	[1624]				22	47	[1127]	17	38	21	39	32	78	[3465]
34	36	[1668]				24	52	[1220]	19	44	24	45			
36	38	[3126]				(25)	(54)	[1253]	20	46	27	54			
38	41	[3490]				26	55	[1282]	23	51	30	56			
(39)	(42)	[3558]				27	56	[1319]	27	61	31	58			
						28	58	[1360]	28	63	33	63			
						29	60	[1411]	29	64	34	65			
						30	62	[1423]	32	71	38	76			
						(31)	(65)	[1480]	34	75	41	80			
						32	68	[1577]	35	76					
						35	79	[3026]	36	77					
						36	81	[3126]	37	78					
						38	83	[3490]	39	82					

The deviation between frequencies of similar modes is always lower than 10 cm^{-1}, except for the modes in parentheses for which the deviation is between 10 and 20 cm^{-1}. Based on the reliable experimental results of A and C and theoretically observed similarities of normal vibrational modes of A and C with the other nucleic acid species, the numbers in brackets represent estimated frequencies (cm^{-1}) for IR and Raman measurements on G, T, AT, and GC (refer to text).

pyrimidines. Most similarities come from ring vibrations and fewer from external atoms to the ring. Modes in column A = G are evenly distributed along the spectrum, allowing for a proper characterization of the purine family.

Vibrations of A or G not appearing in the A = G column represent modes that identify in a unique form the A or G compound in the purine family. The same situation applies to T and C in which the spectrum of T differs almost completely from that of C. Thus, there is practically no need to have recourse to Table V to discriminate T from C in the pyrimidine family.

A main advantage of recognizing similar vibrations is also attributed to the possible extrapolation of results that one may carry from one molecule to another. As an illustration, we consider mode 30 of A and 32 of G. They are similar modes and appear in Table V. In the case of A we have experiments that determine the frequency of mode 30. From Table I their average value (excluding the last measurement to which we assigned the normal mode) is 1423 cm^{-1}. In the case of G measurements involve forces from the environment. Then for mode 32 of G we are able to estimate a value of 1423 cm^{-1} in its gas-phase state. A comparison of this last result with the experimental value of mode 32 in Table III shows a deviation of 64 cm^{-1}. The transferability of values can also be applied to the remaining similar modes between A and G. Then it is possible to quantify the magnitude and direction of environmental forces (in the harmonic approximation) that lead to nonnegligible effects in the vibrational spectra of polycrystalline guanine. The same process may be performed for the C and T modes. The careful portability of results from one species to another, based on similarities between vibrations, allows the calibration of experiments on analogous compounds. As an example of this application, the transferability of experimental results from monomeric species to pair complexes has been performed and results appear

in columns A = AT and C = CG of Table V. We discuss such results in the following sections.

Vibrational Spectra of Pair Complexes

ADENINE-THYMINE

We study the larger Watson–Crick pairs where nucleic acids are the main components. However, the analysis of pairs involves hydrogen bridges. This kind of links represents a different type of bonding than that observed in single NABs. Thus, intermolecular vibrations and perturbations of intramolecular modes brought on by H bridges must be examined carefully.

Limited experimental data exist about pairs[46] in the gas phase, so we perform additional calculations with a different exchange-correlation expression. We use functionals Becke-3 (B3) for exchange and Perdew–Wang for correlation (PW).[47,48] They perform well for energetic features of H-bonded clusters.[49] Contrary to the former BP functional, B3PW is a hybrid expression that contains an explicit mixing of HF exchange. Self-consistent field calculations are done as implemented in the Gaussian 94 package,[35] and the 6–311G orbital basis set and similar cutoffs are as previously discussed. In Gaussian 94 this level of theory is stated as B3PW91/6–311G. In order to be consistent with the monomers, the discussion of dimers is mainly based on the BP results; but attention is also paid to vibrations predicted with the hybrid functional, especially these involving H bridges. The comparison of calculations with different functionals permits us to determine deviations in the frame of DFT and to more reliably quantify changes associated with H links.

The AT compound (Fig. 1) leads to 84 vibrations with 27 out of the plane (Table VI). In the very low frequency region we compute 10 vibrations. Monomers A and T only contribute with a total of 4 modes (refer to Tables I, II). The first modes of dimer AT are different than these found in monomers and appear in a range where experiments in the solid state may face lattice vibrations. Mode 1 describes a butterfly-like oscillation where A and T play the role of wings, while modes 2 and 4 show torsions along the H-bridge interaction region that keep monomers with minimum deformation. Mode 4, for example, simulates an alternating stair. This mode also shows rocking of the CH_3 radical. Mode 2 may be considered a sort of combination of modes 1 and 4. On the other hand, modes 3 and 6–7 represent in-plane bending of one monomer with respect to the other, again along the interaction region with minimum deformation of both fragments. For instance, mode 6 looks like sliding one monomer with respect to the other and mode 7 yields an in-plane approach between A and T. These movements are important because they represent translation motions from the viewpoint of single species as a result of the weak bonding between fragments. The remaining modes, 5 and 8–10, are identified with modes 1–3 and 1 of monomeric T and A.

A good agreement is found between BP and B3PW assignments, but B3PW frequencies are greater than the corresponding BP ones. (Modes 3 and 7 are exceptions in the whole spectrum.) Nevertheless, deviations in the very low frequency region are small and give an average value of 3.6 cm^{-1}, and mode 10 is the largest with 9 cm^{-1}.

The low wave number region presents a pattern characterized in general terms by deformations of ring structures and H-bond waggings. Some vibrations are similar to these computed for monomers and others represent exclusive motions of the dimer. In this last case the modes of fragments are changed because the presence of complementary partners. One of the simplest assignments that characterizes the AT compound is mode 23, which describes the wagging of bond N10—H11 of A, keeping the N10—H12 bond almost static due to the presence of T, or mode 39 with opposite behavior. Other fingerprints of AT are those where atoms of A and T vibrate in a coupled fashion. The most evident examples are modes 12 and 15 that describe the asymmetric and symmetric motions of CH_3 and NH_2. Because coupled oscillations tend to duplicate the number of vibrations observed in monomers, they can also be considered as distinctive factors to chose between the complex and its monomers.

On the other hand, many vibrations of monomers are recognized in the dimer, for instance, those 24 modes of A shown in Table V (column A = AT) and 24 modes of T (column T = AT). In AT there are cases where A shows a larger vibration in comparison with T, but it is the vibration of T that yields a mode similar to the monomeric species and vice versa. Frequency differences between similar vibrations usually represent a small percentage of the frequency value itself. Additional vibrations of monomers, which are different than those that appear in Table V, are also identified in the dimer but they exhibit larger frequency shifts than our cutoff value of 20 cm^{-1}. For exam-

TABLE VI.
Vibrational Spectrum of Watson–Crick Adenine-Thymine Complex.

No.	BP	B3PW	Plane	Assignment
1	27 (6.2)	27 (6.5)	Out	Butt molecule
2	38 (0.2)	38 (0.3)	Out	Tors molecule
3	68 (4.0)	67 (4.0)	In	Bend molecule
4	82 (1.1)	83 (1.4)	Out	Tors molecule
5	112 (0.1)	117 (0.2)	Out	Tors [T], rock CH_3
6	114 (0.8)	115 (0.6)	In	Bend molecule
7	122 (1.9)	119 (1.4)	In	[A] and [T] approaching each other
8	131 (1.8)	136 (1.8)	Out	Tors [T], rock CH_3
9	169 (2.8)	177 (2.9)	Out	Tors [A], wag C6 — NH_2
10	190 (1.7)	199 (2.2)	Out	Tors ring [T] (wag group C2 — N3 — C4)
11	228 (4.2)	237 (4.4)	Out	Butt [A]
12	284 (2.7)	291 (2.1)	In	A bend C5 — CH_3, C6 — NH_2
13	298 (0.1)	313 (0.3)	Out	Tors [A]
14	306 (2.4)	318 (2.7)	Out	Tors [T], wag C5 — CH_3
15	313 (40.2)	319 (42.4)	In	S bend C5 — CH_3, C6 — NH_2
16	389 (76.9)	404 (77.2)	In	A bend C2 — O, C4 — O
17	421 (4.2)	438 (3.2)	Out	Def ring [T], wag C6 — H
18	466 (16.0)	482 (17.7)	In	Sqz ring [T]
19	516 (6.0)	535 (5.2)	In	Bend [A] (sqz group N3 — C4 — N9), bend C6 — NH_2
20	524 (3.5)	543 (6.6)	In	Def R6 [A] (S sqz group C2 — N3 — C4, N1 — C6 — C5)
21	537 (45.4)	560 (35.4)	In	Def ring [T] (sqz group C2 — N3 — C4), A bend N1 — H, C2 — O
22	557 (8.9)	580 (17.5)	Out	Def R6 [A] (wag group N1 — C2 — N3), S wag C2 — H, N10 — H11
23	583 (98.6)	607 (87.3)	Out	Wag N10 — H11
24	594 (22.9)	617 (17.2)	In	S bend C2 — O, C4 — O, C5 — CH_3
25	624 (21.5)	646 (18.8)	In	Def R6 [A]
26	631 (79.7)	655 (82.8)	Out	Wag N1 — H
27	636 (88.9)	666 (70.6)	Out	Def [A] (wag group C5 — N7 — C8), S wag C8 — H, N9 — H
28	651 (45.1)	677 (64.3)	Out	A wag C8 — H, N9 — H
29	666 (20.0)	703 (32.2)	Out	Def R5, R6 [A] (wag group N3 — C4 — N9), wag N9 — H
30	695 (2.8)	723 (3.0)	In	Breath [A]
31	719 (43.8)	756 (21.0)	Out	Def ring [T] (wag group N1 — C2 — N3), wag N1 — H
32	727 (44.8)	756 (36.1)	In	Breath [T]
33	730 (15.5)	771 (51.2)	Out	Def ring [T] (S wag group N1 — C2 — N3, N3 — C4 — C5)
34	775 (19.6)	822 (9.5)	Out	Def R5, R6 [A] (wag group C4 — C5 — C6), wag C8 — H
35	785 (12.9)	817 (13.0)	In	Def ring [T] (sqz group C6 — N1 — C2), str C5 — CH_3
36	799 (1.3)	851 (6.6)	Out	Def R5, R6 [A] (wag group C4 — C5 — C6), wag C8 — H
37	875 (10.1)	908 (12.0)	In	Def R5, R6 [A] (sqz group N1 — C2 — N3)
38	880 (35.2)	929 (8.8)	Out	Def ring [T], wag C6 — H
39	884 (109.2)	896 (166.8)	Out	Wag N10 — H12
40	917 (14.3)	954 (14.9)	In	Def R5 [A] (sqz group N7 — C8 — N9)
41	946 (14.7)	996 (12.1)	Out	Wag C2 — H
42	953 (32.8)	996 (21.2)	In	Def ring [T] (str N1 — C2), S bend C9 — H11, C9 — H10, C9 — H12
43	1004 (2.3)	1043 (4.1)	In	S bend N10 — H11, N10 — H12, C9 — H10, C9 — H11, C9 — H12
44	1008 (42.4)	1047 (38.3)	Out	Def R6 [T], [A], S bend N10 — H11, N10 — H12, S wag C9 — H11, C9 — H10, C9 — H12
45	1052 (24.6)	1099 (27.9)	In	Def R5 (str C8 — N9), A bend C8 — H, N9 — H
46	1056 (2.6)	1095 (2.1)	Out	Def ring [T], wag C9 — H10, C9 — H11, C9 — H12
47	1110 (4.4)	1158 (5.7)	In	Def R5, R6 [A], bend C2 — H, C8 — H, N9 — H
48	1154 (10.5)	1198 (19.6)	In	Def ring [T], A bend N1 — H, C6 — H
49	1161 (94.9)	1176 (115.3)	Out	Wag N3 — H
50	1204 (49.4)	1274 (12.4)	In	Def ring [T] (str C2 — N3, N3 — C4), bend C6 — H, N10 — H11, N10 — H12
51	1208 (50.3)	1253 (82.0)	In	Def ring [T] (str N1 — C6), str C5 — CH_3
52	1214 (19.7)	1265 (32.9)	In	Def R5, R6 [A], S bend C8 — H, N9 — H
53	1232 (39.0)	1282 (35.9)	In	Def R5 [A] (str C5 — N7), S bend N10 — H11, N10 — H12
54	1275 (106.8)	1336 (142.1)	In	Def R5, R6 [A] (str C2 — N3, C5 — N7), A bend C8 — H, N9 — H
55	1291 (80.8)	1349 (42.5)	In	Def R5, R6 [A] (str N1 — C2, C5 — N7), bend C8 — H
56	1335 (2.0)	1388 (10.3)	In	Def R5, R6 [A] (str C8 — N9), bend C2 — H, C8 — H, N9 — H, N10 — H11
57	1335 (37.8)	1392 (26.8)	In	Bend C6 — H
58	1357 (26.4)	1423 (25.0)	In	Def R5, R6 [A] (str C4 — N9, N7 — C8), bend all H [A]
59	1378 (56.5)	1442 (85.5)	In	Def ring [T] (str C4 — C5), S bend N1 — H, C6 — H

TABLE VI.
(Continued)

No.	BP	B3PW	Plane	Assignment
60	1394 (1.0)	1453 (0.3)	In	A bend C2 — H, N9 — H
61	1413 (10.9)	1460 (9.7)	In	Umb CH_3
62	1427 (14.8)	1497 (16.5)	In	Def R5, R6 [A] (str N1 — C6, N7 — C8), S bend all H [A]
63	1445 (34.2)	1508 (46.6)	In	Def ring [T] (str N1 — C6, C4 — C5), bend N1 — H
64	1465 (10.3)	1510 (10.8)	In	Sciss CH_3
65	1467 (0.7)	1534 (21.7)	In	Def R5, R6 [A], str C6 — NH_2, bend C2 — H, C8 — H, sciss C9 — H10, C9 — H12
66	1480 (63.3)	1528 (10.2)	In	Sciss C9 — H10, C9 — H12
67	1499 (194.9)	1561 (260.7)	In	Sciss CH_3, str C4 — O, S bend N1 — H, N3 — H
68	1547 (88.5)	1623 (74.3)	In	Def R5, R6 [A] (str N3 — C4), bend N3 — H, N9 — H, N10 — H11
69	1598 (344.1)	1661 (388.2)	In	Def R6 [A] (str C5 — C6), S bend N3 — H, N10 — H12
70	1622 (8.1)	1689 (1.6)	In	Def R6 [A] (str C5 — C6), def ring [T] (str C5 — C6), A bend N3 — H, N10 — H12
71	1639 (91.7)	1711 (145.4)	In	Def ring [T] (str C4 — C5, C5 — C6), S bend N1 — H, C6 — H
72	1675 (312.3)	1731 (363.5)	In	Sciss NH_2, str C6 — NH_2
73	1717 (316.7)	1784 (481.7)	In	Bend N3 — H, sciss NH_2
74	2414 (3218.4)	2661 (2799.2)	In	Str N3 — H
75	2948 (31.5)	3033 (27.5)	In	S str C9 — H10, C9 — H11, C9 — H12
76	3008 (15.4)	3099 (14.2)	In	A str C9 — H10, C9 — H12
77	3034 (24.1)	3121 (20.3)	In	A str C9 — H11, C9 — H10, C9 — H12
78	3120 (23.0)	3215 (11.8)	In	Str C6 — H
79	3129 (6.5)	3225 (2.7)	In	Str C2 — H
80	3159 (1514.8)	3345 (1214.0)	In	Str N10 — H12
81	3195 (1.6)	3286 (0.7)	In	Str C8 — H
82	3538 (62.1)	3662 (81.4)	In	Str N1 — H
83	3566 (54.9)	3695 (76.4)	In	Str N9 — H
84	3577 (84.8)	3698 (104.4)	In	Str N10 — H11

See footnotes of Table I and Table II for abbreviations. [A], adenine; [T], thymine. They are avoided whenever possible as, for instance, in mode 26 where N1 — H is understood to belong to thymine even when the N1 of adenine induces a hydrogen bridge with H14 (Fig. 1). BP and B3PW indicate the exchange-correlation functionals (levels of density functional theory) used in this work; refer to the text for details.

ple, mode 6 of T at 365 cm^{-1} is recorded at 389 cm^{-1} in AT or mode 19 of A at 917 cm^{-1} is found at 946 cm^{-1} in AT. Such modes are not present in Table V.

The B3PW and BP frequencies differ by an average of 29 cm^{-1}, and mode 36 gives the largest difference of 52 cm^{-1}. At this stage we are unable to find which of the functionals is more accurate.

The intermediate frequency region shows deformations of rings with bending of H bonds. We observe simple vibrations (modes 61, 64 and 72) that correspond to umbrella and scissoring motions of CH_3 and NH_2. Negligible frequency shifts are found for the CH_3 modes 61 and 64 of AT when we compare them with modes 27 and 29 of T. In contrast, mode 72 shows a large shift of 41 cm^{-1} with regard to mode 34 of A. This mode indicates the important role played by the NH_2 group in constructing the AT complex.

Differences between BP and B3PW frequencies appear to increase systematically because they become larger than before. An average deviation of 60 cm^{-1} is found and mode 68 shows the largest difference of 76 cm^{-1}.

The spectrum tail has 11 H stretches and T contributes with 6 and A with 5. (These are the numbers of H stretches in monomers.) Large changes are observed in the complex due to the proximity of fragments. Because A links T to produce bridges N3—H14—N1 and N10—H12—O8 (Fig. 1), the amino group of A and the N3—H bond of T must show large spectral changes. Amino hydrogens that are equivalent in adenine at this interval, now vibrate with different frequencies in the complex. The stretching of N3—H in T shows a great change (1088 cm^{-1}) in going from the single species to the dimer. This bond also takes part in a wagging motion of T (mode 12) and AT (mode 49). The corresponding change is 465 cm^{-1}, which is induced by the formation of H bridges between nucleic fragments.

In the high frequency region the average deviation between the BP and B3PW schemes becomes 123 cm^{-1}. Mode 74 that participates in one of the

H bridges yields the largest deviation of 247 cm^{-1}. In order to assess the accuracy of both approaches, we compare BP and B3PW frequencies with our IR and Raman estimated frequencies for AT given in Table V. Such frequencies were assigned to AT by making use of modes similar to A. In this frequency interval three modes appear and none of them involves an H link. Estimated IR and Raman frequencies for the AT modes 79, 81, and 83 are 3026, 3126, and 3490 cm^{-1}, respectively. The BP scheme gives an average deviation of 83 cm^{-1}, and mode 79 gives the largest value of 103 cm^{-1}. On the other hand, the B3PW scheme gives an average deviation of 188 cm^{-1} and the largest deviation is 205 cm^{-1} (mode 83). If the analysis includes all estimated values that appear in column A = AT of Table V, absolute mean deviations of 25 and 60 cm^{-1} for the BP and B3PW schemes, respectively, are obtained. Except for modes 9 and 13, all B3PW frequencies are higher than the estimated measurements. We deduce that BP frequencies are more accurate than the B3PW ones but assignments are similar in most situations. Differences in the two DFT formulations are attributed to the use of different exchange-correlation functionals and, in particular, to the exchange term because it renders a larger contribution than the correlation term. Note that no H links were included and harmonic frequencies were assumed. Thus, anharmonic effects should be taken into account for a proper evaluation of the BP and B3PW schemes (especially for H bonds where our conclusions may change). However, because of the systematic overestimation of frequencies by the B3PW functional, we believe that B3 largely inherits the behavior of HF exchange and demands the introduction of scaling factors. In the AT compound the scaling factor becomes approximately 0.95. (A similar value for different systems is found if one uses the B3 exchange term in combination with a different correlation functional.[50]) On the other hand, the BP approximation is expected to yield frequencies for H bridges that are in better agreement with the experiment. Nevertheless, the BP and B3PW schemes should still be used with care.

In the BP and B3PW cases vibrations 74 and 80 yield the largest intensities. They refer to stretches of H bonds that take part in the formation of H bridges.

Finally, frequency intervals become densely populated in the AT spectrum that are due to similar bonds in A and T. For large strands of DNA and RNA the population becomes denser. Hence, the characteristic vibrations of single NABs, together with those of dimers, should help to simplify the molecular identification experimentally.

GUANINE-CYTOSINE

Previous results at the HF level have been given for this compound,[51] and the introduction of a scale factor indirectly accounts for electron correlation. The GC complex (Fig. 1) consists of 81 normal modes (Table VII).

The first vibrations in the very low frequency region are of similar nature to those of AT. We summarize them as follows. Mode 1 shows butterfly motion with C and G that simulates wings. Mode 2 is an out of plane torsion of the molecule along the H-bridge interaction region. Mode 3 gives an alternating stair. Modes 4 and 6 look like sliding one fragment with respect to the other, and mode 7 shows an approaching movement of both compounds along lines determined by the three H links. The remaining modes (5 and 8–10) show substantial deformation of single fragments and, with the exception of mode 8, they are not characteristic modes of the complex because such vibrations are identified in single species.

A good agreement is found between BP and B3PW assignments. Their frequencies show an average deviation of 5 cm^{-1}, and mode 10 gives the largest value of 10 cm^{-1}.

In the low frequency region we find a vibrational pattern similar to all previously discussed species: deformation of smaller entities like ring structures, the wagging of atom groups, and a few other modes with negligible deformation of rings. The breathing movements of molecules G and C (modes 24 and 34) yield frequency shifts of 20 cm^{-1} with respect to isolated fragments. The structural symmetry observed in the interaction region between G and C explains the occurrence of some similar modes in the two compounds. For example, modes 14 and 15 are the same but occur at different sites or, equivalently, modes 20 and 22. Hydrogens taking part in the interaction region lead to characteristic vibrations of the GC pair. For example, modes 38, 43, and 46 are also found in single species but with different frequencies. Deviations between B3PW and BP wave numbers are small and have an average value of 29 cm^{-1}. Mode 40 shows the largest deviation of 53 cm^{-1}, and B3PW values are always larger than BP ones.

In the intermediate frequency interval the deformations are due to stretches of the C—N kind and later of the C—C kind. In contrast to the counterparts A and T, the C and G partners contain not

TABLE VII.
Vibrational Spectrum of Watson–Crick Guanine-Cytosine Complex.

No.	BP	B3PW	Plane	Assignment
1	29 (11.6)	30 (12.8)	Out	Butt molecule
2	44 (2.9)	45 (3.0)	Out	Tors molecule
3	85 (0.3)	87 (0.5)	Out	Tors molecule
4	105 (1.1)	103 (1.1)	In	Bend molecule
5	143 (0.0)	150 (0.0)	Out	Tors [G], wag C2 — NH_2
6	145 (34.2)	144 (35.3)	In	Bend molecule
7	150 (2.2)	148 (1.6)	In	[G] and [C] approaching each other
8	167 (1.6)	177 (2.3)	Out	Tors ring [C], A wag C2 — O, C4 — NH_2
9	185 (2.0)	194 (1.8)	Out	Tors [G], wag C6 — O
10	200 (0.6)	210 (0.5)	Out	Def butt [G], tors [C], S wag C2 — O, C4 — NH_2
11	225 (9.3)	235 (10.2)	Out	Def [G] (wag group C2 — N1 — C6), def ring [C]
12	339 (17.8)	348 (17.6)	In	Def [G] (sqz group N3 — C4 — N9), S bend C2 — NH_2, C6 — O
13	347 (1.4)	364 (2.2)	Out	Def [G] (wag group C2 — N3 — C4)
14	370 (5.5)	380 (6.8)	In	A bend C2 — NH_2, C6 — O
15	408 (17.2)	417 (15.2)	In	A bend C4 — NH_2, C2 — O
16	423 (3.6)	444 (3.1)	Out	Def ring [C], wag C6 — H
17	488 (5.6)	505 (5.7)	In	Def R6 [G] (S sqz group N1 — C6 — C5, C2 — N3 — C4)
18	522 (1.6)	543 (2.0)	In	Def R6 [G] (sqz group C2 — N1 — C6), bend N11 — H13
19	534 (52.5)	553 (49.2)	In	Def ring [C], S bend C2 — O, C4 — NH_2
20	540 (94.1)	566 (88.9)	Out	Wag N11 — H13
21	545 (7.0)	563 (7.4)	In	Def [C] (sqz group N3 — C4 — N7), A bend C2 — O, C6 — H
22	553 (66.4)	581 (66.0)	Out	Wag N7 — H10
23	578 (3.7)	598 (4.3)	In	Def ring [C] (sqz group C2 — N3 — C4)
24	621 (10.6)	645 (9.9)	In	Breath [G]
25	622 (136.8)	658 (97.9)	Out	Def R5, R6 [G], S wag C8 — H, N9 — H
26	636 (14.5)	663 (59.8)	Out	Def R5, R6 [G] (wag group N7 — C8 — N9), A wag C8 — H, N9 — H
27	648 (24.2)	678 (27.0)	Out	Def ring [C], wag N1 — H13
28	659 (0.3)	698 (5.3)	Out	Def R5, R6 [G] (wag group N3 — C4 — N9), wag all H [G]
29	669 (46.4)	695 (47.4)	In	Def R5 [G] (sqz group C5 — C4 — N9), S bend C2 — NH_2, C6 — O
30	695 (7.1)	734 (14.2)	Out	Def R5, R6 [G] (wag group N1 — C2 — N3)
31	698 (70.0)	726 (65.3)	Out	Def ring [C] (wag group N3 — C4 — C5), S wag N1 — H13, C5 — H
32	727 (66.8)	774 (67.1)	Out	Def ring [C] (wag group N1 — C2 — N3), wag N1 — H13
33	746 (31.2)	796 (16.8)	Out	Def R5, R6 [G] (wag group N1 — C6 — C5), wag C8 — H
34	756 (0.8)	788 (0.9)	In	Breath [C]
35	773 (18.8)	822 (26.2)	Out	Def R5, R6 [G] (wag group C4 — C5 — C6), wag C8 — H
36	774 (41.0)	813 (55.7)	Out	Def ring [C] (wag group N3 — C4 — C5), S wag C5 — H, C6 — H
37	795 (17.4)	830 (18.3)	In	Def R5, R6 [G] (sqz group C4 — N9 — C8), bend N11 — H13
38	808 (138.8)	835 (165.2)	Out	Wag N11 — H12
39	914 (14.3)	953 (14.4)	In	Def R5 [G] (sqz group N7 — C8 — N9)
40	935 (0.2)	987 (0.2)	Out	A wag C5 — H, C6 — H
41	938 (0.2)	978 (0.1)	In	Def ring [C] (str N1 — C2), S bend C6 — H, N7 — H9, N7 — H10
42	970 (2.4)	1007 (1.7)	In	Def ring [C] (str C4 — C5), S bend C6 — H, N7 — H9, N7 — H10
43	998 (18.8)	1020 (27.7)	Out	A wag N1 — H14, N7 — H9
44	1010 (12.7)	1055 (8.5)	In	Def R6 [G] (str N1 — C2, N3 — C2), S bend N9 — H, N11 — H13
45	1036 (29.4)	1083 (32.5)	In	Def R5 [G] (str C8 — N9), A bend C8 — H, N9 — H
46	1048 (238.5)	1075 (282.9)	Out	S wag N1 — H14, N7 — H9
47	1079 (1.1)	1124 (1.6)	In	S bend N11 — H12, N11 — H13, A bend C5 — H, N7 — H10
48	1085 (53.0)	1127 (47.8)	In	A bend N1 — H13, C5 — H, S bend N7 — H9, N7 — H10
49	1127 (36.3)	1170 (23.1)	In	S bend C5 — H, N7 — H9, N7 — H10

TABLE VII.
(Continued)

No.	BP	B3PW	Plane	Assignment
50	1140 (24.0)	1187 (6.1)	In	S bend C8 — H, N9 — H, S bend N11 — H12, N11 — H13
51	1141 (17.3)	1199 (39.0)	In	Def R6 [G] (str N1 — C6), N11 — H13
52	1221 (50.3)	1264 (56.1)	In	Bend all H [C]
53	1244 (82.0)	1310 (76.6)	In	Def ring [C] (str N1 — C2, C2 — N3), bend N7 — H10
54	1257 (1.3)	1312 (7.5)	In	Def R5 [G] (str N7 — C8), bend C8 — H
55	1305 (69.9	1371 (71.8)	In	Def R5, R6 [G] (S str C4 — N9, C5 — N7)
56	1329 (15.2)	1386 (9.1)	In	Def R5 [G] (str C8 — N9), A bend C8 — H, N9 — H
57	1361 (6.5)	1410 (10.8)	In	S bend N1 — H13, C5 — H, C6 — H
58	1373 (10.0)	1435 (11.7)	In	Def R5, R6 [G] (A str N1 — C2, C4 — C5), S bend N9 — H, N11 — H13
59	1392 (78.9)	1450 (105.2)	In	Def ring [C] (str N1 — C6, N3 — C4), bend all H [C]
60	1420 (22.1)	1480 (100.2)	In	Def R5, R6 [G], str C2 — NH_2, bend N1 — H14
61	1438 (123.8)	1506 (85.6)	In	Def R5, R6 [G] (S str C2 — N3, N7 — C8), str C6 — O, bend N1 — H14
62	1470 (51.8)	1537 (53.8)	In	Def R6 [C], [G], bend N1 — H13, N1 — H14, N7 — H10, N11 — H13
63	1480 (21.4)	1547 (25.5)	In	Def R5, R6 [G], [C] (str C4 — C5), bend N11 — H13, all H [C]
64	1515 (112.8)	1572 (130.1)	In	Def ring [C], str C4 — NH_2, S bend N1 — H13, C5 — H, C6 — H
65	1550 (152.3)	1620 (155.0)	In	Def R5, R6 [G] (str N3 — C4, C4 — N9), bend N9 — H
66	1602 (570.9)	1671 (604.9)	In	Def R6 [C], [G], str C2 — O, bend N1 — H13, N1 — H14
67	1615 (488.7)	1684 (494.4)	In	Def R6 [C], [G] (str C5 — C6), sciss NH_2 [C], A bend C5 — H, C6 — H
68	1633 (6.9)	1701 (40.6)	In	Def R6 [C], [G] (str C5 — C6), str C2 — O, C6 — O, bend N1 — H13
69	1676 (685.5)	1734 (616.4)	In	Sciss NH_2 [G], str C2 — NH_2, bend N1 — H14
70	1689 (207.4)	1749 (475.0)	In	Sciss NH_2 [G] and [C], bend N1 — H14
71	1717 (89.4)	1775 (228.7)	In	Sciss NH_2 [C], bend N1 — H14
72	2804 (299.1)	3017 (59.8)	In	S str N1 — H14, N7 — H9
73	2888 (3210.6)	3081 (2976.3)	In	A str N1 — H14, N7 — H9
74	3128 (4.2)	3222 (2.5)	In	A str C5 — H, C6 — H
75	3151 (10.4)	3245 (4.2)	In	S str C5 — H, C6 — H
76	3192 (4.9)	3284 (2.0)	In	str C8 — H
77	3203 (1458.0)	3355 (1294.6)	In	Str N11 — H12
78	3527 (59.3)	3650 (77.5)	In	Str N1 — H13
79	3564 (54.9)	3688 (71.0)	In	Str N7 — H10
80	3565 (33.0)	3696 (55.1)	In	Str N9 — H
81	3598 (80.5)	3718 (98.3)	In	Str N11 — H13

See footnotes of Table I for abbreviations. [G], guanine; [C], cytosine. They are avoided whenever possible as, for instance, in mode 78 where N1 — H13 is understood to belong to cytosine (Fig. 1). BP and B3PW indicate the exchange-correlation functionals (levels of density functional theory) used in this work; refer to the text for details.

only similar NH_2 groups but also similar C—O bonds (Fig. 1). This explains the presence of many scissoring motions and C—O and C—NH_2 stretches in the last part of this region. Assignments obtained with the B3PW and BP schemes are essentially the same, but BP frequencies are systematically smaller than the B3PW. (The mean deviation is 59 cm^{-1}, and mode 65 gives the largest difference of 71 cm^{-1}.) This result is due to the design of exchange-correlation expressions, which is mainly based on energetics rather than on vibrational aspects.[52]

The last part of the GC spectrum refers to 10 H stretches. Only five are characteristic of the complex and modes 72, 73, and 77 involve at least one H bridge. The largest intensities are computed for such vibrations and they also present the largest deviations between BP and B3PW frequencies along the spectrum. If we compare BP and B3PW frequencies with estimated IR and Raman values from Table V (column C = GC), B3PW yields a mean deviation of 53 cm^{-1} while BP gives 50 cm^{-1}. No H bridges are considered there, but the comparison process includes values along the whole spectrum. This fact calls attention to the possible overemphasis of the HF component in the B3 hybrid exchange. B3PW frequencies require a scale factor of 0.98 to match our estimated IR and Raman frequencies, while in AT it was 0.95. Thus, functional B3 requires several scale factors, de-

pending on the compound and the frequency interval[53] under study.

Many normal modes of the G and C monomers are found in the complex (Table V). The number of similarities indicates the coupling strength between fragments that build the pair. In the GC pair we find several normal modes where atoms of G and C simultaneously participate in vibrations of the complex. Such vibrations are due to three H bridges that confer higher rigidity to the GC compound and also explain why GC shows a lower number of similar normal modes in comparison with AT.

The normal modes with largest intensities are 73, 77, and 69. They all involve H bridges and represent markers of GC.

Note that the analysis of the GC compound, as well as for the previous nucleic acids, is by no means complete and it represents an overview of the vibrational features.[54] Undoubtedly, details proliferate for the large number and complexity of normal modes, but the information contained in the tables should complement our discussion.

Binding Energies

Binding energies (defined as the total energy of the molecule minus the energies of all atoms taken independently) represent key features to investigate the stability of NABs and become of interest when studying mutation processes. Table VIII shows that purines are more stable than pyrimidines (G is followed by A, T, and C). In the case of pairs, GC is less stable than AT. However, the H bridge energy for the AT dimer [computed as $E(\mathrm{AT}) - [E(\mathrm{A}) + E(\mathrm{T})]$, where $E(X)$ is the binding energy of species X] is just -0.030 au in comparison with the -0.054 au of GC that contains an additional link. The average energy per H bridge is -0.015 and -0.018 au in the AT and GC pairs, respectively. Thus, the H bridges in GC look stronger than in AT.

Zero-point vibrational energies of NABs (see Table VIII) diminish the strength of molecular binding and shift binding energies upward because they have similar values. The 6–311G basis set, initially optimized for HF calculations, gives energies in good agreement with these of ref. 38 that were obtained with basis sets of double-zeta quality. In fact, our binding energies show an almost constant shift of 0.1 au in comparison with these of ref. 38, but the zero-point energies are practically the same.

Finally, dipole moments and the A, B, and C rotational constants for the NABs and their pairs are considered in Table VIII. Discrepancies emerge among our dipoles and those from ref. 38. The G and GC compounds yield the largest deviations of approximately 0.50 debyes (D). Basis sets that include polarization functions are expected to give results in better agreement with the experiment.[55,56] In the case of rotational constants, which may be considered as additional identifiers of tautomeric species,[56–59] basis sets play a less important role in the DFT calculations. Our results and

TABLE VIII.
Binding and Zero-Point Vibrational Energies (au), Dipole Moments (D) and Rotational Constants (MHz) of DNA Nucleic Acids and Their Watson–Crick Pairs.

Species	Energy	ZPVE	Dipole	Rotational Constants
A	−2.657	0.109	2.49 [2.58]	2311.58, 1527.46, 919.72 (2371.87, 1573.36, 946.26)
T	−2.608	0.112	4.54 [4.44], (4.13)	3128.71, 1362.56, 954.80 (3201.21, 1404.81, 982.61)
G	−2.825	0.113	7.25 [6.75, 6.50]	1862.02, 1086.70, 686.21 [1910.48, 1112.24, 703.44]
C	−2.227	0.096	6.90 [6.56], (6.0–6.5)	3764.46, 1960.56, 1289.16 (3871.55, 2024.93, 1330.31)
AT	−5.295	0.222	1.61 [1.75]	885.93, 181.47, 150.76
GC	−5.106	0.212	6.68 [6.12]	858.21, 200.63, 162.62

Binding energies do not include the zero-point vibrational energy (ZPVE) correction. Numbers in parentheses represent experimental measurements and in brackets represent calculated values with a different level of density functional theory than the one used in this work. For the dipole moments refer to refs. 38, 55, and 56 and for the rotational constants refer to refs. 56–59.

these given in ref. 56 show small deviations and they are close to experimental data. For the complex pair systems, we expect that rotational constants show the same level of accuracy.

Conclusions

We calculated the harmonic vibrational spectra of nucleic acids in the frame of DFT. The Kohn–Sham version was applied with gradient corrections in exchange and correlation energy terms. The results were in good agreement with IR and Raman measurements with average deviations of 5–10%. All nucleic acids followed a similar pattern; at very low frequencies in-plane and out of plane torsion and bending of molecules were observed. In the low frequency region, five- and six-membered rings showed a diversity of in-plane and out of plane deformations that was due to angle squeezing and wagging. They appeared together with wagging of H bonds. At intermediate frequencies in-plane deformations of ring units were the most common movements. They were associated with angle squeezing and C—N and C—C stretches and also showed bending of H bonds. At the tail of the spectra only H stretches were observed. The vibrations of NH_2 and CH_3 groups were documented along the spectra.

A similar analysis was performed for Watson–Crick pairs. However, two different exchange-correlation energy functionals were used (BP and B3PW) to more reliably quantify changes associated with H bridges. A similar vibrational trend was found as in the monomers. Many normal modes of single species were recognized in the dimers, allowing one to establish by complementarity the vibrational fingerprints of Watson–Crick pairs. The BP frequencies were closer to our estimated experimental frequencies than the B3PW ones. The B3PW frequencies appeared larger than the BP values, and their differences increased systematically from the low to the high frequency region. The largest deviations were obtained for H bridges, indicating that for H-bonded systems the use of the exchange-correlation functional becomes critical.

Binding, zero-point, and hydrogen-bond energies; dipole moments; and rotational constants were also calculated to analyze the stability of NABs and provide additional parameters for their better identification. With the exception of dipole moments, the orbital basis sets used here generally gave good results compared with the experiment.

The vibrational spectra may be of help in genetics for the identification of nucleic acids in biocompounds. They might be taken as reference data to determine the magnitude and direction of environmental forces, which are capable of spectral manifestation, and in the parameterization of force fields.

References

1. Herzberg, G. Infrared and Raman Spectra of Polyatomic Molecules; Van Nostrand Reinhold: New York, 1945.
2. Stepanian, S. G.; Sheina, G. G.; Radchenko, E. D.; Blagoi, Y. P. J Mol Struct 1985, 131, 333.
3. Hirakawa, A. Y.; Okada, H.; Sasagawa, S.; Tsuboi, M. Spectrochim Acta 1985, 41A, 209.
4. Dhaouadi, Z.; Ghomi, M.; Austin, J. C.; Girling, R. B.; Hester, R. E.; Mojzes, P.; Chinsky, L.; Turpin, P. Y.; Coulombeau, C.; Jobic, H.; Tomkinson, J. J Phys Chem 1993, 97, 1074.
5. Majoube, M. J Mol Struct 1986, 143, 427.
6. Majoube, M. J. Chim Phys 1984, 81, 303.
7. Sheina, G. G.; Stepanian, S. G.; Radchenko, E. D.; Blagoi, Y. P. J Mol Struct 1987, 158, 275.
8. Szczesniak, M.; Szczepaniak, K.; Kwiatkowski, J. S.; KuBulat, K.; Person, W. B. J Am Chem Soc 1988, 110, 8319.
9. Yarkony, D. R., Ed. Modern Electronic Structure Theory; World Scientific: River Edge, NJ, 1995; Vols. 1 and 2.
10. Del Bene, J. E. J Phys Chem 1983, 87, 367.
11. Florián, J. J Mol Struct (Theochem) 1992, 253, 83.
12. Hrouda, V.; Florián, J.; Hobza, P. J Phys Chem 1993, 97, 1542.
13. Leszczynski, J. Chem Phys Lett 1990, 174, 347.
14. Leszczynski, J. Int J Quantum Chem Quantum Biol Symp 1992, 19, 43.
15. Leszczynski, J. J Mol Struct (Theochem) 1994, 311, 37.
16. Wiórkiewicz–Kuczera, J.; Karplus, M. J Am Chem Soc 1990, 112, 5324.
17. Riggs, N. V. Chem Phys Lett 1991, 177, 447.
18. Gould, I. R.; Hillier, I. H. Chem Phys Lett 1989, 161, 185.
19. Estrin, D. A.; Paglieri, L.; Corongiu, G. J Phys Chem 1994, 98, 5653.
20. Kwiatkowski, J. S.; Leszczynski, J. J Phys Chem 1996, 100, 941.
21. Paglieri, L.; Corongiu, G.; Estrin, D. A. Int J Quantum Chem 1995, 56, 615.
22. Florián, J.; Leszczynski, J.; Scheiner, S. Mol Phys 1995, 84, 469.
23. Donnamaria, M. C.; Maranon, J.; Howard, E. I.; Fantoni, A.; Grigera, J. R. Mol Simul 1996, 18, 101.
24. (a) Levine, I. N. Quantum Chemistry; Prentice Hall: Englewood Cliffs, NJ, 1991; (b) Richards, W. G.; Cooper, D. L. Ab Initio Molecular Orbital Calculations for Chemists; Oxford University Press: Oxford, UK, 1985.

25. Grev, R. S.; Janssen, C. L.; Schaefer, H. F. III. J Chem Phys 1991, 95, 5128.
26. Hohenberg, P.; Kohn, W. Phys Rev 1964, B136, 864.
27. Kohn, W.; Sham, L. J. Phys Rev 1965, A140, 1133.
28. Labanowski, J. K.; Andzelm, J. W., Eds. Density Functional Methods in Chemistry; Springer–Verlag: New York, 1991.
29. Seminario, J. M.; Politzer, P., Eds. Modern Density Functional Theory: A Tool for Chemistry; Elsevier Science: New York, 1995.
30. Szabo, A.; Ostlund, N. S. Modern Quantum Chemistry; McGraw–Hill: New York, 1989.
31. Fulde, P. Electron Correlation in Molecules and Solids, Springer Series in Solid-State Sciences No. 100; Springer–Verlag: New York, 1993.
32. Parr, R. G.; Yang, W. Density Functional Theory of Atoms and Molecules, International Series of Monographs on Chemistry No. 16; Oxford University Press: Oxford, UK, 1989.
33. (a) Becke, A. D. Phys Rev 1988, A38, 3098; (b) Perdew, J. P. Phys Rev 1986, B33, 8822 [Erratum Phys Rev 1986, B34, 7406].
34. In spite of giving large differences in energy, the LSD and BP gradient-corrected approximations yield frequencies and assignments close to each other for single nucleic acids. For H-bonded pairs differences between the two schemes emerge, and the LSD scheme requires improvement.
35. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head–Gordon, M.; Gonzalez, C.; Pople, J. A. Gaussian 94; Gaussian Inc.: Pittsburgh, PA, 1995.
36. Broo, A.; Holmen, A. Chem Phys 1996, 211, 147.
37. Schlegel, H. B. J Comput Chem 1982, 3, 214.
38. Santamaria, R.; Vázquez, A. J Comput Chem 1994, 15, 981.
39. (a) Santamaria, R.; Quiróz–Gutiérrez, A.; Juárez, C. J Mol Struct (Theochem) 1995, 357, 161; (b) Santamaria, R.; Cocho, G.; Corona, L.; González, E. Chem Phys 1998, 227, 317.
40. (a) Johnson, B. G.; Gill, P. M. W.; Pople, J. A. J Chem Phys 1993, 98, 5612; (b) Fan, L.; Ziegler, T. J Chem Phys 1992, 96, 9005.
41. (a) Qin, Y.; Wheeler, R. A. J Chem Phys 1995, 102, 1689; (b) Bérces, A.; Ziegler, T. J Chem Phys 1993, 98, 4793.
42. (a) Kieninger, M.; Suhai, S. Int J Quantum Chem 1994, 52, 465; (b) Bérces, A.; Ziegler, T. Chem Phys Lett 1993, 203, 592; (c) Kaschner, R.; Seifert, G. Int J Quantum Chem 1994, 52, 957; (d) Sim, F.; St-Amant, A.; Papai, I.; Salahub, D. R. J Am Chem Soc 1992, 114, 4391.
43. (a) Fournier, R.; Papai, I. In Recent Advances in Density Functional Methods; Chong, D., Ed.; World Scientific: Singapore, 1996; Vol. 1; (b) Bérces, A.; Ziegler, T. J Phys Chem 1994, 98, 13233; (c) Bérces, A.; Ziegler, T.; Fan, L. J Phys Chem 1994, 98, 1584.
44. Barone, V.; Adamo, C.; Russo, N. Int J Quantum Chem 1994, 52, 963.
45. In the absence of nuclear symmetry, a total of $3N$-6 vibrations are obtained for each compound, where N is the number of atoms. With regard to the nomenclature in this work, a bending (wagging) movement is understood as an in-plane (out of plane) vibration.
46. Florián, J.; Leszczynski, J.; Johnson, B. G. J Mol Struct 1995, 349, 421.
47. Becke, A. D. J Chem Phys 1993, 98, 5648.
48. Perdew, J. P.; Wang, Y. Phys Rev 1992, B45, 13244.
49. Hobza, P.; Sponer, J.; Reschel, T. J Comput Chem 1995, 16, 1315.
50. Urban, J.; Schreiner, P. R.; Vacek, G.; von Rague Schleyer, P.; Huang, J. Q.; Leszczynski, J. Chem Phys Lett 1997, 264, 441.
51. Florián, J.; Leszczynski, J. Int J Quantum Chem 1995, S22, 207.
52. (a) Santamaria, R. Int J Quantum Chem 1997, 61, 891; (b) Santamaria, R.; Kaplan, I. G.; Novaro, O. Int J Quantum Chem 1995, 56, 307, 421.
53. Florián, J.; Baumruk, V.; Leszczynski, J. J Phys Chem 1996, 100, 5578, and references therein.
54. Supplementary material available from the authors like ground-state geometries and other electronic properties may be requested by e-mail. The visualization of normal modes and animation of vibrations are expected to appear soon in our web page.
55. Kulakowska, I.; Geller, M.; Lesyng, B.; Wierzchowski, K. L. Biochim Biophys Acta 1974, 361, 119.
56. Nowak, M.; Lapinski, L.; Kwiatkowski, J. S.; Leszczynski, J. Computational Chemistry: Reviews of Current Trends; Leszczynski, J., Ed.; World Scientific: River Edge, NJ, 1997; Vol. 2 and references therein.
57. Brown, R. D.; Godfrey, P. D.; McNaughton, D.; Pierlot, A. P. Chem Phys Lett 1989, 156, 61.
58. Brown, R. D.; Godfrey, P. D.; McNaughton, D.; Pierlot, A. P. J Am Chem Soc 1989, 111, 2308.
59. Brown, R. D.; Godfrey, P. D.; McNaughton, D.; Pierlot, A. P. JCS Chem Commun 1989, 37.